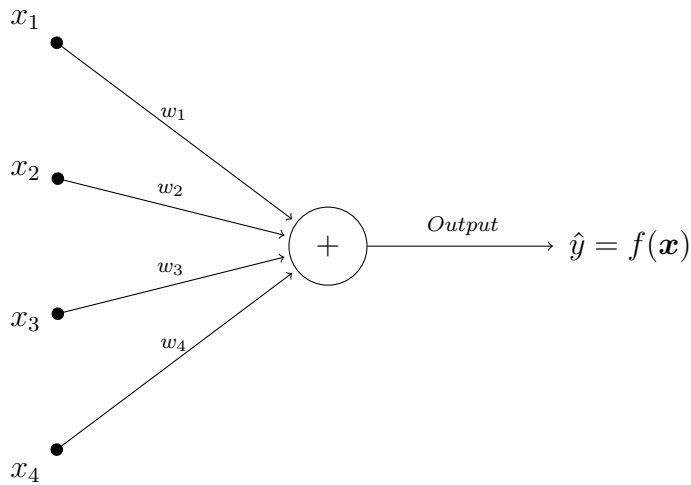<u>Machine Learning</u>

1. Structure

$$\boldsymbol{x} = \begin{Bmatrix} x_1 \\ \vdots \\ x_D \end{Bmatrix} \boxed{\text{Data Structure}} \qquad\qquad \boldsymbol{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_D \end{bmatrix}$$

$\boxed{\text{Machine Structure}}$

$$\hat{y} = f(\boldsymbol{x}) = \boldsymbol{w}^\top \boldsymbol{x} + b \tag{1}$$

$$= \sum_{i=1}^{D} w_i x_i + b \tag{2}$$

2. Criterion = optimize **w** and b (e.g. minimize the mean square error)

$$\underset{\boldsymbol{w},b}{\text{minimize}} \; \underset{p(x,y)}{\mathbb{E}} \left\{ |y - (\boldsymbol{w}^\top \boldsymbol{x} + b)|^2 \right\}$$

↗

Expectation w.r.t. density
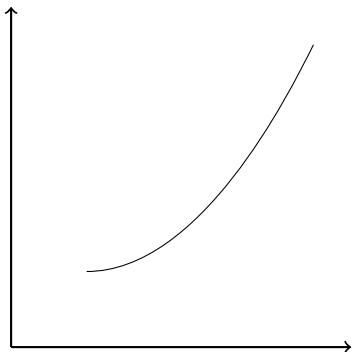
3. Algorithm

4. Primal *arrow* Dual